Comments of Science Review Panel on the Michigan Water Assessment Tool

Meeting of the Groundwater Conservation Advisory Council

East Lansing, Michigan, 11-12 December 2006

Panel Members: Hal Beecher, Joe DePinto, LeRoy Poff, Bill Woessner

The Science Review Panel views the overall goals and approach used by the Groundwater conservation Advisory Council to develop a science based screening tool as progressive, and we encourage their completion of this effort. We are impressed with the conceptual approach and scientific rigor used to develop linked groundwater-surface-water-ecological models that form the basis of the assessment tool. The amount and quality of empirical data being incorporated into model development are impressive, and they provide a strong basis for developing and implementing a science-based, statewide decision tool that can inform sustainable water and fisheries management of Michigan streams and groundwater.

Based on our review of the documents provided, hearing the presentations made at the meeting, and discussing specific elements and details with the participants, we conclude that the process being developed by the Groundwater Conservation Advisory Council is on sound footing. Because our review was (appropriately) conducted at a point in time when our feedback could be incorporated into final stages of the process, our comments are mostly focused on the need to sharpen the scientific rationale of specific components of the screening tool and to more clearly describe the process to the general public. We do not foresee any serious impediment to achieving screening tool implementation, recognizing that the tool will be refined as additional data and analyses become available.

Here, we present our thoughts on several elements of the proposed tool.

Model Landscape Characterization:

<u>Spatial grain</u>. Modeling the effects of water withdrawal (whether groundwater or surface water pumping) on surface water flows and fish habitat can be accomplished over a range of spatial scales. The landscape characterization criteria used to establish a spatial grain for the assessment tool needs more justification. Based on available data, the finest grain could be at the scale of stream segments (i.e., stream lengths between tributary nodes in a drainage network). Currently, there are about 30,000 "arcs" or stream segments in the database. Coarser grains are possible and should be evaluated, e.g., Seelbach and Wiley's "valley segment" approach lumps the 30,000 reaches into about 15,000 segments. The Council should give more consideration to the grain at which they can meaningfully characterize the entire state landscape. The criteria for making this decision should reflect a tradeoff between describing the scale of the key processes incorporated in the tool (spatial extent of groundwater pumping on the stream flow, scale of variation in index flow,

appropriate spatial scale for characterizing a fish assemblage) and the availability of data necessary to "calibrate" the landscape characterization. The ease of use during implementation should also be a consideration.

Continuous vs. Aggregated Representation.
Separately from the grain of the landscape characterization, it is important to decide how to represent the landscape in a modeling framework. One option is to do so in a "continuous" fashion, that is by representing each of the 30,000 spatial elements by its position in a drainage network. In this approach, a 1^{st}-order stream would flow into a 2^{nd}-order (or higher) stream into a 3^{rd}-order (or higher) stream, etc. The advantage of this approach is spatial coherence among landscape units; this would allow for routing of flows, etc. through a stream network. This seems to be the current approach. A second and perhaps more useful approach would be to *classify* the 30,000 (or fewer aggregate) stream segments according to a number of environmental characteristics, e.g., streamflow and temperature characteristics, landcover types, watershed features, etc. Such classes might include settings such as "cold, small streams", "large, mainstem rivers", etc. This approach would create a smaller number of "classes" of stream settings could be, perhaps, more readily used for statewide management purposes for the assessment tool, because geographically-separated stream segments having similar attributes could be identified. It is also possible that there would be a class-specific approach for estimation of allowable withdrawal.

A clear advantage of the classification approach is that it reduces the complexity of the landscape composition, from the ca. 30,000 individual reaches to some smaller number of stream-type classes that have similar features defined in terms of defining environmental features, such as low flow yield, temperature, and possibly others (e.g., nutrient loads, geology). Transforming the data structure from 30,000 individual reaches to a smaller number of stream classes may possibly liberalize (slightly) the amount of water that can be taken out, because minor stream reaches will be added together resulting in defined segments containing larger flows per zone. A strong advantage of this approach, however, would be that it allows for assessment of multiple withdrawals with a given area (class comprised of several similar stream segments). By contrast, a possible disadvantage could be that highly sensitive headwater reaches might be subsumed in a coarser aggregation, and some thought should be given to the issue of whether they require special consideration.

Irrespective of which landscape characterization is adopted for the statewide screening tool, we explicitly support the concept of making this a dynamic process by keeping track of cumulative depletions over time as additional wells or withdrawal pipes are installed. This updated information should be used to evaluate allowable streamflow depletions computed by the tool.

Streamflow Model.
The approach described appears generally sound; however, some of the specific justifications for the approach need to be clarified. For example, why is an *index flow* selected to be the median monthly flow for the low flow month? We assume this low flow level is viewed as an ecological bottleneck with both direct and indirect impacts on the fish community, but no material to support this rationale was presented to us. Previous research that underpins this criterion needs to be explicitly referenced.

We believe more diagnostics are needed for the regression model that relates low flow discharges to various environmental variables. While the regression model performs reasonably well, there are several outliers that should be examined more systematically to address the question of whether the model is missing some important independent variable(s). For example, could human population density be an important predictor variable? Can non-linear variables and interaction terms be included in the regression? Or, is there a stratification variable that might be used to split the state into a priori regions and thus allow development of region-specific regression models? Or, are relatively large errors associated with absolute values of streamflow (i.e., small deviations in predicted values for small streams can translate into large relative errors)? We recognize that any regression model will have unexplained variation and prediction error, so it is important to state when the performance of a regression model is adequate, given the project goals. Beyond this, it is relevant to consider how sensitive the regression model might be to longer term variation in flow records, as might be associated with changes in land use, natural climatic variation, or human-caused climate change.

We do have some concerns about the lack of precision in developing flow exceedence curves for each of 30,000 arcs. Based on data you have worked with, how much uncertainty/variance is there in these estimates? It strikes us that this approach far exceeds the spatial grain of the calibrating data (i.e., the USGS gauges). Again, a coarser grain of landscape characterization and a classification framework would probably be more amenable to a regionalization approach for developing flow duration curves for classes of stream, as these classes are more likely to have a gauge associated with them. Further, it may be useful to incorporate "snapshot" data into the database to assist in calibration of the index flow estimates for ungauged stream segments. Apparently, there are hundreds to thousands of single-observation measurements of stream discharge recorded throughout the state, and these should be evaluated for their information content and incorporated, as appropriate, in your analyses and calibration process. We note that it may be desirable at some future time to collect additional such point measurements to refine the model estimates of low flow yield.

The network analysis was presented as a second part of the streamflow analysis to ensure that flow in a network proceeds in a logical additive fashion downstream. This is a useful aspect of the flow tool.

Groundwater Model.
As with landscape considerations, we suggest that the representation of groundwater systems be grouped by geologic setting or hydrogeologic setting. This would lump similar settings based on common characteristics. We would suggest that you also need to consider developing vertical geologic setting models for typical conditions in these geologic settings. The hydrogeologic setting approach would allow evaluation of a wider range of conditions than those that can be addressed by the currently proposed analytical solution methods. It is possible that proposed analytical models will give reasonable estimates; however, this should be evaluated by developing simple 2D and 3D numerical box models that are complex enough to represent generalized conditions representative at the state level. If the analytical solutions are not appropriate, then these few (5 or 6?) numerical box models may be useful in generating a range of solutions that can be generalized and applied in implementation of the tool.

In addition, the issue of the timing of stream flow depletion in response to initiation of groundwater pumping needs to be evaluated and resolved. Such time dependent depletion is an important part of evaluating the response of the stream to pumping.

Finally, an attempt at characterizing the uncertainty of stream depletion modeling is needed. This factor may be partly derived by running a number of simulations spanning a reasonable range of parameter values and comparing and contrasting results. It could also include developing different reasonable conceptual models and examining the sensitivity of models to parameter variations. This should be done not on all models but a select group representing a range of conditions.

Ecological Model.
At a statewide scale, ecological complexity can be overwhelming. The identification of "assemblage types" and the dominant constituent species of each assemblage type is a well-reasoned and widely-accepted way to proceed in deriving useful ecological metrics for broad geographic application. Michigan is fortunate to have high quality and spatially extensive fisheries data, and this information enables the development of defensible metrics for use in the assessment tool.

The general approach being undertaken is good. A small number of fish assemblage types can be identified across the state based on natural associations of species that are distributed under a wide range of conditions in key environmental factors (low flow yield, temperature). However, the statistical support for these assemblage types needs to be more clearly presented. The proposal to develop "habitat suitability indices" (HSI) for the dominant fish species for each assemblage type is reasonable, but it may be useful to identify which of the "dominant" species for each assemblage type is most sensitive to a unit change in low flow discharge and to use this species to define the "habitat factor". Further, it may be desirable to define the HSI in terms of other key environmental variables besides low flow values, such as temperature and nutrient status. Some effort should be directed at determining how sensitive fish assemblage types are to changes in key environmental variables, i.e., how much of a change in flow/temperature conditions is required to cause a "shift" in fish assemblage type? In this respect, it would also be useful to examine assemblage sensitivity with respect not only to changes in abundance of the "dominant" species but also to changes in the abundances of the other species comprising the assemblage. For some assemblage types, it may be possible to also employ population-based models that can predict population density as a function of temperature or low flow habitat. This would be most profitably explored for assemblages characterized by game species, for which much may be known.

"Sensibility Analysis" of Screening Tool.

When this tool is completely developed (beta phase), there will be substantial scientific uncertainty that reflects the propagation of uncertainty through each of the linked component models. This cannot be avoided and does not invalidate the implementation of the tool. If the policy goal of tool implementation is to identify those proposed pumping projects that have a high likelihood of impairment of stream function, then the tool must somehow be "conservatively" applied (given the underlying scientific uncertainty). The proper "balance" would presumably evolve over time; however, it would be informative and useful to know how the tool might perform prior to its implementation.

We recommend that the sensitivity of the assessment tool be tested by examining a wide range of potential withdrawal scenarios under varied hydrogeologic, stream segment and fish assemblage conditions that are representative throughout the state. The range of scenarios to be tested should span a variety of determinants, such as distance of well from stream, depth of well, withdrawal rates and schedules, different stream categories in terms of hydrogeologic conditions and representative fish assemblages. This would provide information on how the tool might respond to these potential scenarios in terms of screening them in or out, and thus achieving the desirable "balance" in the projected implementation of the tool. This diagnostic sensitivity analysis should be conducted before finalizing some of the recommended thresholds and parameter decisions in the tool. Also, the tool should be evaluated by comparing its predicted results with those in areas of the state where detailed river-fish-groundwater evaluations have been completed (e.g. the Muskegon watershed contains both data and a process model calibrated to these data). This will provide an additional means by which the tool can be evaluated and its uncertainty estimated.

We believe that in the long-term the state should move toward development of a linked hydrologic-water-quality-ecological process-oriented model for each of the state's watersheds, starting with those with the highest potential stress from withdrawal applications.